View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

ACADEMIC PRESS
Available online at www.sciencedirect.com
SCIENCE DIRECT®
Advances in Mathematics 177 (2003) 280–296
ADVANCES IN Mathematics
http://www.elsevier.com/locate/aim

Time-delay coordinates and polynomial mappings

Mark Pollicott

Department of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Received 28 March 2001; accepted 29 April 2002
Communicated by R.D. Mauldin

Abstract

The use of time-delay coordinates to reconstruct mappings is well known and provides an important practical tool in studying real-world problems. In this note we formulate the underlying mathematical analysis in the natural context of polynomial mappings and real analytic systems. This is particularly well adapted to systems defined by simple algebraic equations where, unlike in the general case, we do not require techniques from differential geometry.
© 2003 Elsevier Science (USA). All rights reserved.

MSC: 58F

Keywords: Time delay; Embedding; Dynamical systems; Polynomial maps

0. Introduction

Many experiments give as their output long records of numbers produced by the observed value of measurable quantity at regular time intervals. The most frequently used method for the reconstruction of the phase space of a dynamical system is the method of time-delay coordinates. Pakard et al. [6] originally demonstrated this principle with the reconstruction of the solution to the Lorenz attractor.

A particularly simple model of a dynamical system is a diffeomorphism $f: M \to M$ of a compact finite-dimensional manifold M. The manifold M typically represents a state space for the dynamical system. The "experimental data" can be taken to be a

E-mail address: mp@maths.man.ac.uk.

0001-8708/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0001-8708(02)00018-X

finite family of C^∞ function $g : M \to \mathbb{R}$ and the sequence of scalers

$$g(x), g(fx), g(f^2x), \ldots, g(f^nx), \ldots \tag{0.1}$$

arising from evolution from an initial state x (where, as usual, f^n represents the n-fold composition of f with itself). The method of time-delay coordinates involves reconstructing M by an embedding $\Phi : M \to \mathbb{R}^N$, for some appropriate N and under suitable hypotheses on g, defined by $\Phi(x) = (f(x), g(fx), \ldots, g(f^{N-1}x)) \in \mathbb{R}^N$ (cf. [1,2] for a general overview).

Both Takens [9] and Sauer et al. [7] gave a rigorous basis for the method in the case of generic systems. More precisely, the theorems of Takens and Sauer, Yorke and Casdagli show rigorously that for typical *differentiable* n-dimensional manifolds M and diffeomorphisms f the manifold can be realized as an embedding into $\mathbb{R}^{2n+1}$ and the diffeomorphism $f : M \to M$ can be reconstructed from data (1) using the mapping Φ. The difference of viewpoint is that Takens result is categorical and the result of Sauer, Yorke and Casdagli deals with the more physical notion of prevalence. As Kolmogorov observed [3]: "*Typical behaviour approach from the categorical side is interesting more as a tool of proving existence results .... while an approach from the measure theoretic side seems to be more physically reasonable and natural.*"

We take as our starting point that many of the physical systems actually studied can be either described or approximated by polynomial or algebraic equations. This setting seems particularly appropriate in light of the development of algorithmic and computer-oriented theories (cf. [10]) where working with polynomial equations is more natural. Furthermore, a not inconsiderable technical advantage is that the ambient spaces (of polynomials) are finite dimensional, and so much of the differential geometry machinery used in the original approaches is reduced to elementary linear algebra.

A preliminary problem we describe is the question of when N polynomial mappings $g_1, \ldots, g_N : M \to \mathbb{R}$ have the property that the map $g : M \to \mathbb{R}^N$ defined by $g(x) = (g_1(x), \ldots, g_N(x))$ is an embedding. This question is also discussed in [7].

In Sections 1 and 2 we recall the usefulness of real algebraic approximations to more general smooth systems. In Section 3 we consider the analogue of the results of Sauer, Yorke and Casdagli in the corresponding setting of polynomial mappings. For completeness, we illustrate the use of time-delay coordinates in Section 4, by the trivial example of a rotation of the plane and the more complicated Henón map of the plane. In the appendices we give the proofs of our main results in Section 3.

1. Real algebraic manifolds and Nash's theorem

In this section we recall some classical results that demonstrate that although we are dealing with real algebraic systems, these can always be used to closely model differentiable systems.

Assume that we have a diffeomorphism $f: M \to M$ on an abstract compact C^∞ manifold, of dimension $n \geqslant 2$. Without loss of generality, we can always consider an n-dimensional manifold M as a submanifold of the Euclidean space $\mathbb{R}^N$, for sufficiently large N, by the following classical result of Whitney.

Whitney's Theorem. *Every C^∞ n-dimensional compact manifold M is diffeomorphic to an n-dimensional real analytic submanifold $M' \subset \mathbb{R}^N$ (i.e., there exists a diffeomorphism $f: M \to M'$ and M' has an atlas of coordinate charts which are real analytic)* [11].

It is well known that in the Whitney's theorem we can always take $N = 2n + 1$. Below we shall state the extension of this result by Nash gives that this embedded manifold can be even be taken to be a real analytic variety. Let us recall the definition.

Definition. A real algebraic variety V can also be described as the zero set of a finite set of real polynomials $f_1, \dots, f_k$, say, i.e., $V = \{x \in \mathbb{R}^N: f_i(x) = 0, i = 1, \dots, k\}$.

The following elegant result of Nash[1] can be viewed as a refinement of the Whitney theorem.

Nash's Theorem. *Every C^∞ compact manifold M is diffeomorphic to a real algebraic manifold $V \subset \mathbb{R}^N$ (i.e., there exists a diffeomorphism $f: M \to V$)* [5].

The original paper of Nash contains a very readable account of the proof.

We shall be interested in those diffeomorphisms $f: V \to V$ which respect the real algebraic structure. In particular, we shall assume that $f: V \to V$ is the restriction of a polynomial map $f: \mathbb{R}^N \to \mathbb{R}^N$.

2. Transcendence dimension and immersions

Let us begin with some consequences of working in the context of real algebraic manifolds. Following Nash's original paper [5] we can use an equivalent, more algebraic, definition of a real algebraic manifold. Let $\mathscr{R}$ be a ring of functions $g: M \to \mathbb{R}$. We say that a set $\{s_1, \dots, s_k\} \subset \mathscr{R}$ is algebraically independent if whenever we find a rational polynomial P such that $P(s_1, \dots, s_k) = 0$ then P must be identically zero. We define the *transcendence degree* d of $\mathscr{R}$ to be the maximal cardinality k of such sets.

An equivalent definition of a *real algebraic manifold* is that it consists of a real analytic manifold M and a ring $\mathscr{R}$ of functions $g: M \to \mathbb{R}$ such that:

(1) each $f \in \mathscr{R}$ is a real analytic function;
(2) the *transcendence degree* of $\mathscr{R}$ is n, i.e., given any functions $g_1, \dots, g_{n+1} \in \mathscr{R}$ they satisfy a non-trivial polynomial equation; and

[1] Nash famously received the Nobel prize for economics for his early work in game theory [4].

(3) there is a finite set $f_1, \dots, f_N : M \to \mathbb{R}$ of *basic functions* such that $x \mapsto (f_1(x), \dots, f_N(x)) \in \mathbb{R}^N$ are an embedding.

To illustrate the definition, we can consider a trivial example.

Trivial Example. We can consider the two-dimensional sphere

$$S^2 = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}.$$

We consider the ring of functions $\mathscr{R}$ which are polynomials restricted to M. The projections $f_i(x) = x_i, i = 1, 2, 3$ give rise to the identity map $x \mapsto (f_1(x), f_2(x), f_3(x))$ embedding the sphere. The ring $\mathscr{R}$ has transcendence dimension 2, e.g., the functions $f_1(x), f_2(x), f_3(x)$ satisfy the polynomial equation $f_1(x)^2 + f_2(x)^2 + f_3(x)^2 = 1$.

Using a more familiar geometric perspective, we can let $T_x V$ denote the tangent space to V at a point $x \in V$. The *dimension* of the variety V corresponds to the value d for which $d = \dim T_x V$ on an open dense set and $\dim T_x V \geqslant d$ at all points. (The possibility of $\dim T_x V \geqslant d$ occurs when x is a point of self-intersection of the variety V. When V is actually an embedded manifold, then $d = \dim T_x V$ will hold at all points.)

The algebraic nature of the real analytic manifold V means that many geometric features can be read-off from the ring $\mathscr{R}$. This includes the dimension:

Lemma. *The transcendence dimension of the ring $\mathscr{R}$ is the same as the dimension of the variety V* [5].

We say that a map $g : V \to \mathbb{R}^N$ is an *immersion* if the map $D_x g$ has maximal rank at all points $x \in V$. The map $g : V \to g(V)$ need not be one-to-one, but since $g(V)$ is an algebraic variety its self-intersections must necessarily be subvarieties. This has immediate, if rather simple, consequences.

For definiteness, let $x_1, \dots, x_N : V \to \mathbb{R}$ be the projection onto the coordinates and let $\mathscr{S}_N \subset \mathscr{R}$ be the ring they generate. If the transcendence degree d of the algebra $\mathscr{S}_N$ they generate satisfies $d = n$ then the map

$$g : V \to \mathbb{R}^N,$$

$$g : x \mapsto (x_1(x), x_2(x), \dots, x_N(x))$$

is an immersion. Otherwise, if $d < n$ then g will not only fail be an immersion, but $g(V)$ will have dimension $d < n$.

Similarly, we can consider related time-delay coordinates. Let $x_1 : V \to \mathbb{R}$ be a projection onto one coordinate. Given a rational map $T : V \to V$ we can consider the compositions

$$x_1, x_1 \circ T, x_1 \circ T^2, \dots, x_1 \circ T^{m-1} \quad \text{for some } m \geqslant 0.$$

If the transcendence degree d of the algebra $\mathscr{R}_m \subset \mathscr{R}$ they generate satisfies $d = n$ then the map

$$\Phi : V \to \mathbb{R}^m,$$

$$\Phi : x \mapsto (x_1(x), x_1(Tx), \ldots, x_1(T^{m-1}x))$$

is an immersion. Otherwise, if $d < n$ then Φ will not only not be an immersion, but $g(V)$ will have dimension $d < n$.

3. Prevalent embeddings and approximations

The basic problem we shall consider in this section is to find a realization of an abstract dynamical system $f : M \to M$, including an embedding of M into $\mathbb{R}^{2n+1}$, say, using typical polynomial functions $g : M \to \mathbb{R}^{2n+1}$. We recall that an *embedding* of a manifold M into $\mathbb{R}^{2n+1}$, say, is defined to be a map $g : M \to \mathbb{R}^d$ such that:

(1) g is an immersion; and
(2) g is a one-to-one map onto its image.

We want to formulate prevalence results in the context of real algebraic manifolds. We need to introduce some notation which will be useful in the statement and proof of Theorem 1 below. We want to consider the polynomial map $g : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ given by

$$g(x_1, \ldots, x_{2n+1}) = (g^{(1)}(x_1, \ldots, x_{2n+1}), \ldots, g^{(2n+1)}(x_1, \ldots, x_{2n+1})).$$

For each $(2n+1)$-tuple $\underline{k} = (k_1, \ldots, k_{2n+1})$ of positive integers we write $|\underline{k}| = k_1 + \cdots + k_{2n+1}$. If each of the $2n+1$ polynomial maps $g^{(i)} : V \to \mathbb{R} (i = 1, \ldots, 2n+1)$ is of degree at most d, say, then they can be written in the form

$$g^{(i)}(x_1, \ldots, x_{2n+1}) = \sum_{|\underline{k}| \leqslant d} a_{\underline{k}}^{(i)} x^{\underline{k}},$$

where $x^{\underline{k}} = x_1^{k_1} \ldots x_{2n+1}^{k_{2n+1}}$ denotes a monomial and $a_{\underline{k}}^{(i)} \in \mathbb{R}$ is the corresponding coefficient. Let $N = N(2n+1, d)$ denote the dimension of the space of polynomials in $2n+1$ variables of degree at most d.[2] There is a natural correspondence between $\mathbb{R}^N$ and the space of such polynomials, given by

$$\mathbb{R}^N \ni (a_{\underline{k}}^{(i)})_{|\underline{k}| \leqslant d} \mapsto g^{(i)},$$

[2] These values are easily computed. The dimension of the space polynomials in $2n+1$ variables of degree exactly d is $\binom{2n+d}{d}$.

for each $i = 1, \ldots, 2n+1$. There is a corresponding natural finite-dimensional parameterisation of the $(2n+1)$-tuples of polynomials of degree d by the finite-dimensional Euclidean space $\mathbb{R}^{N+(2n+1)}$, given by

$$\mathbb{R}^{N+(2n+1)} \ni (a_{\underline{k}}^{(i)})_{\substack{|\underline{k}| \leqslant d \\ 1 \leqslant i \leqslant 2n+1}} \mapsto g = (g^{(1)}, \ldots, g^{(2n+1)}).$$

The next theorem shows that for "typical" polynomial maps $g : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ of this form, the restriction to an n-dimensional manifold $M \subset \mathbb{R}^{2n+1}$ is an embedding. Here we mean that the class of atypical polynomials (i.e., those which do not give embeddings) are particularly small in that they correspond to parameters in $\mathbb{R}^{N+(2n+1)}$ lie on (countable unions) of codimension one manifolds.

Theorem 1. *Let $M \subset \mathbb{R}^{2n+1}$ be an n-dimensional manifold. There exists a countable union of codimension one submanifolds $X \subset \mathbb{R}^{N+(2n+1)}$ such that the polynomial map $g : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ is an embedding for M, provided the parameters $(a_{\underline{k}}^{(i)})_{\substack{|\underline{k}| \leqslant d \\ 1 \leqslant i \leqslant 2n+1}}$ corresponding to g do not lie on X.*

The proof appears in Appendix A.

Remark. In particular, Theorem 1 gives us that the exceptional set of parameters $X \subset \mathbb{R}^{N+(2n+1)}$ has measure zero. This coincides with the conclusions of [7].

4. Time delay coordinates and embeddings

We can write a typical polynomial map $g : M \to \mathbb{R}$ in the form

$$g(x_1, \ldots, x_{2n+1}) = \sum_{|\underline{k}| \leqslant d} a_{\underline{k}} x^{\underline{k}}, \tag{4.1}$$

where $(a_{\underline{k}})_{|\underline{k}| \leqslant d} \in \mathbb{R}^N$. There is a natural correspondence between $\mathbb{R}^N$ and the space of such polynomials, given by

$$\mathbb{R}^N \ni (a_{\underline{k}}^{(i)})_{|\underline{k}| \leqslant d} \mapsto g.$$

The method of time-delay coordinates gives an approach to reconstructing the dynamics of the original diffeomorphism from data (0.1). In particular, the following theorem can be viewed as a variant of Theorem 1.

In the interests of clarity of exposition we do not present the most general hypotheses.

Theorem 2. *Let $M \subset \mathbb{R}^{2n+1}$ be an n-dimensional manifold. Let $f : M \to M$ be a diffeomorphisms whose coordinate functions are polynomial valued. Assume that*

$f: M \to M$ does not have any periodic points $\underline{x} \in M$ of period less than $2n+1$. Then providing d is sufficiently large there exists a countable union of codimension one submanifolds $X \subset \mathbb{R}^N$ such that the map $\Phi_{(f,g)}: M \to \mathbb{R}^{2n+1}$ defined by

$$\Phi_{(f,g)}(x) = (g(x), g(fx), \ldots, g(f^{2n}x))$$

is embedding for M whenever the parameters $(a_{\underline{k}})_{|\underline{k}| \leqslant d}$ corresponding to g do not lie on X.

The proof appears in Appendix B.

In particular, Theorem 2 gives us that the exceptional set of parameters $X \subset \mathbb{R}^N$ has measure zero. This coincides with the conclusions of [7].

Remark. Of course, the dynamics of the original map can also be read-off from the embedding. If we denote the image $\bar{M} = \Phi_{(f,g)}(M)$ then $f: M \to M$ is conjugate to the map $\bar{f}: \bar{M} \to \bar{M}$ defined by

$$\bar{f}(g(y), g(fx), \ldots, g(f^{2n})) = (g(fy), g(f^2x), \ldots, g(f^{2n+1}))$$

(i.e., $\Phi_{(f,g)} \circ f = \bar{f} \circ \Phi_{(f,g)}$).

Remark. One can also make the stronger assertion that the embedding map $\Phi_{(f,g)}$ typically preserves certain dimensions of invariant subsets.

5. Examples

We now recall some well-known examples arising from polynomial equations. They serve to give simple illustrations of the types of systems to which the methods of Theorems 1 and 2 apply.

5.1. The delayed logistic map

A particularly simple algebraic system to begin with is the delayed logistic map $f: \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f(x, y) = (y, ay(1-x)).$$

Consider the specific choice $a = 2.15$. The orbit of $(x, y) = (0.25, 0.25)$ is illustrated in Fig. 1(i), where we plot points $(x^{(n)}, y^{(n)}) = f^n(x, y)$ for $n = 1, \ldots, 10^5$.

Figs. 1(ii)–(iv) show reconstructions of the same orbit using some different choices of function g. In particular, Fig. 1(ii) corresponds to the trivial choice $g(x) = x$. Figs. 2(iii) and (iv) correspond to the choices $g(x) = x - x^3$ and $g(x) = x^2$, respectively.

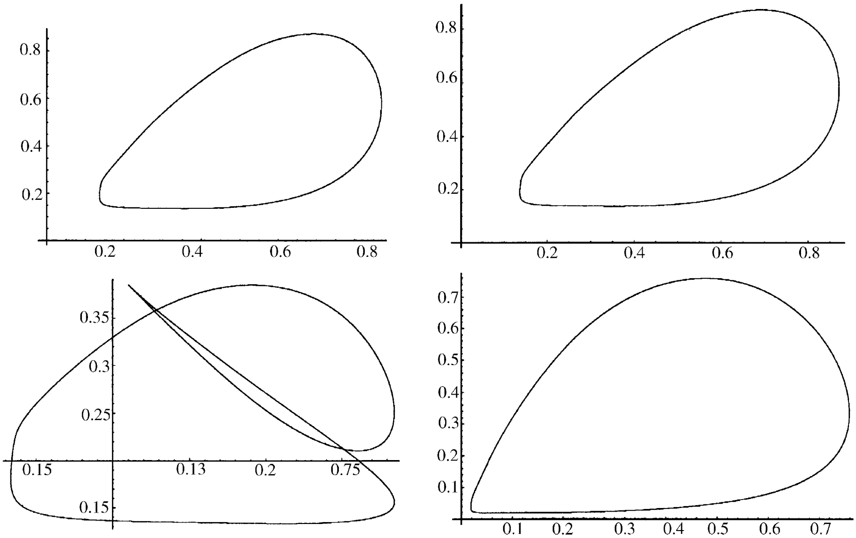

Fig. 1. A plot for the delayed logistic map of: (i) $(x^{(n)}, y^{(n)})$; (ii) $(x^{(n)}, x^{(n+1)})$; (iii) $(g(x^{(n)}), g(x^{(n+1)}))$ for $g(x) = x - x^3$; and (iv) $(g(x^{(n)}), g(x^{(n+1)}))$ for $g(x) = x^2$.

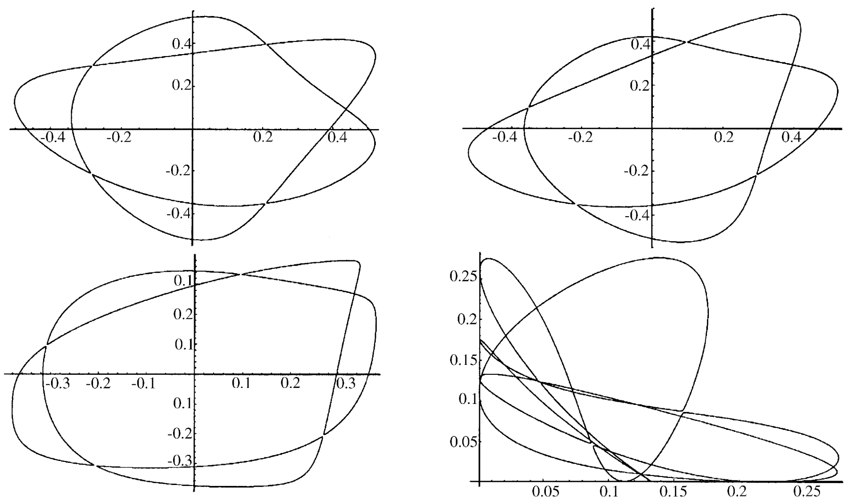

Fig. 2. A plot for the area preserving Henon map of: (i) $(x^{(n)}, y^{(n)})$; (ii) $(x^{(n)}, x^{(n+1)})$; (iii) $(g(x^{(n)}), g(x^{(n+1)}))$ for $g(x) = x - x^3$; and (iv) $(g(x^{(n)}), g(x^{(n+1)}))$ for $g^{(x)} = x^2$.

5.2. *The area preserving Henon map*

An interesting algebraic system is the area preserving Henon map $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f(x, y) = (\cos(\alpha)x + \sin(\alpha)y + \sin(\alpha)x^2, \sin(\alpha)x + \cos(\alpha)y - \cos(\alpha)y^2)$$

for some real parameter α. This can be viewed as a perturbation of the rotation. Let us fix the choice $\alpha = 1.3$. An orbit segment for this map is illustrated in Fig. 2(i), where we plot points $(x^{(n)}, y^{(n)}) = f^n(x, y)$ for $n = 1, \ldots, 10^5$, with the choice of initial point $(x, y) = (0, 0.35325)$.

Figs. 2(ii)–(iv) show reconstructions of the same orbit using the different choices of function g used before.

The next two examples describe discrete polynomial maps given by approximation to well-known flows.

5.3. The Rossler attractor

We can consider the flow on $\mathbb{R}^3$ given by the differential equations

$$\dot{x} = -y - z,$$

$$\dot{y} = x + 0.2y,$$

$$\dot{z} = 0.2 + z(x - 5).$$

We can associate to this flow a discretisation $f : \mathbb{R}^3 \to \mathbb{R}^3$ using the forward Euler method, i.e.,

$$f(x, y, z) = (x + \Delta t(-y - z), y + \Delta t(x + 0.2y), z + \Delta t(0.2 + z(x - 5))),$$

where Δt is a small increment in time. For Δt sufficiently small the orbit segments of f approximate those for the flow.

For definiteness, let us fix $\Delta t = 0.005$. An orbit segment for the discrete map $f : \mathbb{R}^3 \to \mathbb{R}^3$ is illustrated in Fig. 3(i), where we plot points $(x^{(n)}, y^{(n)}, z^{(n)}) = f^n(x, y, z)$ for $n = 1, \ldots, 10^5$, with choice of initial point $(x, y, z) = (-8.0578, 0.6288, 0.0154)$.

Figs. 3(ii)–(iv) show various different reconstructions of the same orbit using time-delayed coordinates and the same functions g we used before.

5.4. The Lorenz attractor

Finally, we consider the flow on $\mathbb{R}^3$ given by the differential equations

$$\dot{x} = 10(y - x),$$

$$\dot{y} = 28x - y - xz,$$

$$\dot{z} = xy - \tfrac{8}{3}z.$$

As in the previous example, we can associate to this flow a discretisation $f : \mathbb{R}^3 \to \mathbb{R}^3$ using the forward Euler method, i.e.,

$$f(x, y, z) = (x + \Delta t(10(y - x)), y + \Delta t(28x - y - xz), z + \Delta t(xy - \tfrac{8}{3}z)),$$

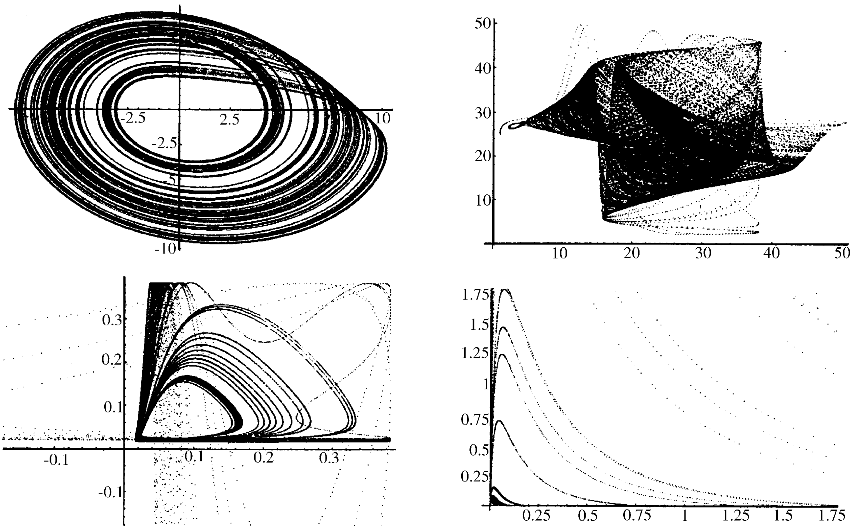

Fig. 3. A plot for the Rossler attractor of: (i)$(x^{(n)}, y^{(n)})$; (ii)$(x^{(n)}, x^{(n+1)})$; (iii) $(g(x^{(n)}), g(x^{(n+1)}))$ for $g(x) = x - x^3$; and (iv) $(g(x^{(n)}), g(x^{(n+1)}))$ for $g(x) = x^2$.

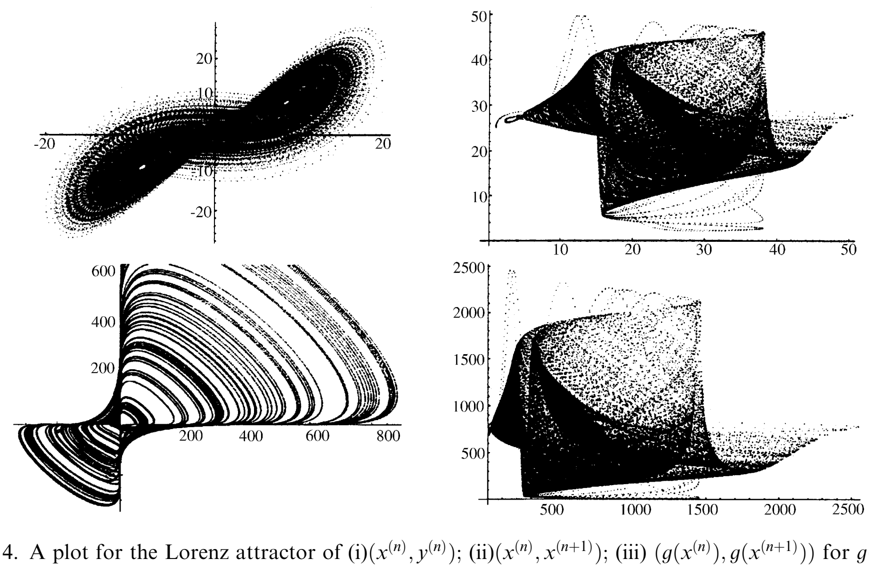

Fig. 4. A plot for the Lorenz attractor of (i)$(x^{(n)}, y^{(n)})$; (ii)$(x^{(n)}, x^{(n+1)})$; (iii) $(g(x^{(n)}), g(x^{(n+1)}))$ for $g(x) = x - x^3$; and (iv) $(g(x^{(n)}), g(x^{(n+1)}))$ for $g(x) = x^2$.

where Δt is a small increment in time. For Δt sufficient small the orbit segments of f approximate those for the flow.

For definiteness, let us fix $\Delta t = 0.005$. In Fig. 4(i), we plot the orbit segment $(x^{(n)}, y^{(n)}) = f^n(x, y)$ for $n = 1, \ldots, 10^5$, with the choice of initial point $(x_0, y_0, z_0) = (1, 1, 1)$.

Figs. 4(ii)–(iv) show various different reconstructions of the same orbit using time-delayed coordinates and the same functions g we used previously.

Acknowledgments

I am grateful to the referee for useful comments and advice on examples.

Appendix A. Proof of Theorem 1

To show that g is one-to-one for typical

$$\underline{a} = (a_{\underline{k}}^{(i)})_{\substack{|\underline{k}| \leqslant d \\ 1 \leqslant i \leqslant 2n+1}} \in \mathbb{R}^{N+(2n+1)}$$

we need to consider the map

$$G : \mathbb{R}^{N+(2n+1)} \times (M \times M - \Delta) \to \mathbb{R}^{2n+1}$$

given by $G(\underline{a}, x, y) = g(x) - g(y)$, where $\underline{\in} \mathbb{R}^{N+(2n+1)}$ and $\Delta = \{(x, x) \in M \times M\}$. The condition for g to be one-to-one is that $G(\underline{a}, M \times M - \Delta) \subset \mathbb{R}^{2n+1}$ does not contain the origin $\underline{0} \in \mathbb{R}^{2n+1}$. We want to show that the set

$$\left\{\underline{a} \in \mathbb{R}^{N+(2n+1)} \colon \underline{0} \notin G(M \times M - \Delta)\right\},$$

i.e., the set of those $\underline{a}$ such that $G(M \times M - \Delta)$ does not contain $\underline{0}$, is contained in a countable union of codimension one manifolds.

We claim that the derivative

$$D_1 G : \mathbb{R}^{N+(2n+1)} \to \mathbb{R}^{2n+1}$$

always has full rank (i.e., rank $2n+1$). If we consider the derivative with respect to just one of the coefficients $a_{\underline{k}}^{(i)}$, $|\underline{k}| \leqslant d$ and $1 \leqslant i \leqslant 2n+1$, then it takes the particular form

$$\partial_{a_{\underline{k}}^{(i)}} G(\cdot, \underline{x}, \underline{y}) = \underbrace{(0, \ldots, 0, x_1^{k_1} \cdots x_{2n+1}^{k_{2n+1}} - y_1^{k_1} \cdots y_{2n+1}^{k_{2n+1}}, 0 \ldots, 0)}_{\text{Only the } i\text{th entry is non-empty}} \in \mathbb{R}^{2n+1},$$

where $\underline{k} = (k_1, \ldots, k_{2n+1})$. In particular, we see that $\partial_{a_{\underline{k}}^{(i)}} G(\cdot, \underline{x}, \underline{y}) \neq 0$ provided $(\underline{x}^{\underline{k}} - \underline{y}^{\underline{k}})_{\underline{k}} \neq 0$, which is automatically true for some $\underline{k}$ since we are assuming $(\underline{x}, \underline{y}) \notin \Delta$. This implies that the derivative

$$D_1 G : \mathbb{R}^{N+(2n+1)} \to \mathbb{R}^{2n+1}$$

always has full rank.

Assume, for a particular parameter value $\underline{b} \in \mathbb{R}^{N+(2n+1)}$, there exists a point $(x, y) \in M \times M - \Delta$ such that $G(x, y) = \underline{0} \in \mathbb{R}^{2n+1}$, then we want to show that locally this condition only holds on a codimension one manifold in $\mathbb{R}^{N+(2n+1)}$. The proof becomes more transparent if we apply the following simple lemma on coordinate changes.

Rank Lemma. *Assume that $k \geqslant m$ and let $U \subset \mathbb{R}^k$ be an open set and $f : U \to \mathbb{R}^m$ a C^1 map such that the derivative $Df(\underline{w})$ has maximal rank m for all $w \in U$. Then we can choose local coordinates with respect to which we may write*

$$f(w_1, \dots, w_m, w_{m+1}, \dots, w_k) = (w_1, \dots, w_m).$$

Proof. This is essentially a corollary of the Inverse Function Theorem [8, p.56]. Let us assume, by permuting coordinate labels, if necessary, that the $m \times m$ matrix given by $(\partial f_i / \partial w_j)_{i,j=1}^m$ has non-zero determinant. We want to choose new coordinates in a neighbourhood of a given by

$$\bar{w}_i(\underline{w}) = \begin{cases} f_i(w_1, \dots, w_k) & \text{for } i = 1, \dots, m, \\ w_i & \text{for } i = m+1, \dots, k. \end{cases}$$

The derivative for this coordinate change $\underline{w} \mapsto \bar{w}$ is $\begin{pmatrix} \frac{\partial f_i}{\partial w_j} & \cdots \\ 0 & I \end{pmatrix}$ which has non-zero determinant, and thus by the implicit function theorem is a diffeomorphism in some neighbourhood. Moreover, by construction we can write $(f_1, \dots, f_m) = (\bar{w}_1, \dots, \bar{w}_m)$, and dropping the overline gives the required result. $\square$

Using charts, we can identify a neighbourhood of $(\underline{x}, \underline{y})$ in $M \times M - \Delta$ with a neighbourhood in $\mathbb{R}^{2n}$. We then want to apply the Rank Lemma in the case of a neighbourhood $U \subset \mathbb{R}^{N+(2n+1)} \times \mathbb{R}^{2n}$ of $(\underline{b}, \underline{x}, \underline{y})$. The set $V \subset \mathbb{R}^{2n+1}$ corresponds to a neighbourhood of $G(\underline{x}, \underline{y})$. Since $D_1 G$ has maximal rank $2n+1$ we conclude that the full Frechet derivative DG also has rank $2n+1$. Using the Rank Lemma it is particularly easy to see that, in a neighbourhood of $\underline{b}$, those $\underline{a}$ such that $G(\underline{a}, \underline{x}, \underline{y}) = 0$ must lie on a codimension one manifold. This shows that g is one-to-one, except for those parameters on a countable union of codimension one manifolds X_0 in $\mathbb{R}^{N+(2n+1)}$.

To show that $g : M \to \mathbb{R}^{2n+1}$ is an immersion for typical $\underline{a} \in \mathbb{R}^{N+(2n+1)}$ we need to consider the map

$$H : \mathbb{R}^{N+(2n+1)} \times SM \to \mathbb{R}^{2n+1}$$

given by $H(\underline{a}, \underline{x}, v) = D_{\underline{x}} g(v)$, where SM is the space of unit tangent vectors to M.

The condition for $g: M \to \mathbb{R}^{2n+1}$ to be an immersion is that $\underline{0} \notin H(\underline{a}, SM)$. Given $(\underline{x}, v) \in SM$ we can write

$$H(\underline{a}, \underline{x}, v) = \begin{pmatrix} \frac{\partial g^{(1)}}{\partial x_1}(\underline{x}) & \frac{\partial g^{(1)}}{\partial x_2}(\underline{x}) & \cdots & \frac{\partial g^{(1)}}{\partial x_{2n+1}}(\underline{x}) \\ \frac{\partial g^{(2)}}{\partial x_1}(\underline{x}) & \frac{\partial g^{(2)}}{\partial x_2}(\underline{x}) & \cdots & \frac{\partial g^{(2)}}{\partial x_{2n+1}}(\underline{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g^{(2n+1)}}{\partial x_1}(\underline{x}) & \frac{\partial g^{(2n+1)}}{\partial x_2}(\underline{x}) & \cdots & \frac{\partial g^{(2n+1)}}{\partial x_{2n+1}}(\underline{x}) \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ \vdots \\ \vdots \\ v_{2n+1} \end{pmatrix}, \qquad \text{(A.1)}$$

where the final vector in (A.1) represents v. We can write out explicitly the partial derivatives which appear as entries in the first matrix in (A.1) by

$$\frac{\partial g^{(i)}}{\partial x_j}(\underline{x}) = \sum_{1 \leqslant |\underline{k}| \leqslant d} (k_j a_{\underline{k}}^{(i)}) \underline{x}^{\underline{k}(j)}, \quad i, j = 1, \ldots, 2n+1, \qquad \text{(A.2)}$$

where $\underline{k}^{(j)} = (k_1, \ldots, k_{j-1}, k_j - 1, k_{j+1}, \ldots, k_{2n+1})$. For the particular choice $\underline{k} = (0, \ldots, 0, 1, 0, \ldots, 0)$, where the entry 1 occurs in the jth place, if we take the partial derivative of the expression for $\partial g^{(i)}/\partial x_j$ in (A.2) with respect to the parameter $a_{\underline{k}}^{(i)}$ then we get

$$\partial_{a_{\underline{k}}^{(i)}} H(\cdot, \underline{x}) = \underbrace{\begin{pmatrix} 0 \ldots 0 \ldots 0 \\ \vdots \ddots \vdots \ddots \vdots \\ 0 \ldots 1 \ldots 0 \\ \vdots \ddots \vdots \ddots \vdots \\ 0 \ldots 0 \ldots 0 \end{pmatrix}}_{\text{Only the } (i,j) \text{ entry is non-zero}}$$

We can deduce that the map $D_1 H : \mathbb{R}^{N+(2n+1)} \times \mathbb{R}^{2n} \to \mathbb{R}^{2n+1}$ has full rank (i.e., rank $2n+1$).

Assume that we start from a choice of $\underline{b}$ such that $H(\underline{b}, \underline{x}, v) = 0$. Using charts, we can identify a neighbourhood of $(\underline{x}, v)$ in SM with a neighbourhood in $\mathbb{R}^{2n-1}$. We then want to apply the Rank Lemma in the case of a neighbourhood $U \subset \mathbb{R}^{N+(2n+1)} \times \mathbb{R}^{2n-1}$ of $(\underline{b}, \underline{x}, v)$. The set $V \subset \mathbb{R}^{2n+1}$ corresponds to a neighbourhood of $H(\underline{x}, v)$. Since $D_1 H$ has maximal rank $2n+1$ we conclude that the full Frechet derivative DH also has rank $2n+1$. Using the Rank Lemma it is particularly easy to see that, in a neighbourhood of $\underline{b}$, those $\underline{a}$ such that $H(\underline{a}, \underline{x}; v) = 0$ must lie on a codimension two manifold. This shows that g is one-to-one, except for those . parameters on a countable union of codimension two manifolds X_1 in $\mathbb{R}^{N+(2n+1)}$.

Finally, if we let $X = X_0 \cup X_1$ then we see that providing $\underline{a} \notin X$ the associated map $g: M \to \mathbb{R}^{2n+1}$ is not an embedding. This completes the proof of Theorem 1. $\square$

Appendix B. Proof of Theorem 2

Proof of Theorem 2. To show that g is one-to-one we would like to know that the derivative

$$D_1 G : \mathbb{R}^{N+(2n+1)} \to \mathbb{R}^{2n+1},$$

where $G(\underline{x}, \underline{y}) = (g(\underline{x}) - g(\underline{y}), g(f\underline{x}) - g(f\underline{y}), \dots, g(f^{2n}\underline{x}) - g(f^{2n}\underline{y}))$, has maximal rank (i.e., rank $2n+1$). In particular, we need to consider the derivatives

$$\partial_{(\underline{a})_{|\underline{k}| \leqslant d}} G(\cdot, \underline{x}, \underline{y}) : \mathbb{R}^N \to \mathbb{R}^{2n+1}$$

with respect to each of the coefficients $\underline{a} \in \mathbb{R}^N$, where $G(\underline{x}, \underline{y}) = \Phi_{(f,g)}(\underline{x}) - \Phi_{(f,g)}(\underline{y})$. These derivatives take the form

$$\begin{aligned}\partial_{\underline{a}} G(\cdot, \underline{x}, \underline{y}) = (&x_1^{k_1} \cdots x_{2n+1}^{k_{2n+1}} - y_1^{k_1} \cdots y_{2n+1}^{k_{2n+1}}, \dots, \\ &[f^{2n+1}(\underline{x})]_1^{k_1} \cdots [f^{2n+1}(\underline{x})]_{2n+1}^{k_{2n+1}} - [f^{2n+1}(\underline{y})]_1^{k_1} \cdots [f^{2n+1}(\underline{y})]_{2n+1}^{k_{2n+1}}),\end{aligned}$$

where $\underline{k} = (k_1, \dots, k_{2n+1})$ and the ith term in $\partial_{\underline{a}} G(\cdot, \underline{x}, \underline{y})$ is of the form

$$[f^i(\underline{x})]_1^{k_1} \cdots [f^i(\underline{x})]_{2n+1}^{k_{2n+1}} - [f^i(\underline{y})]_1^{k_1} \cdots [f^i(\underline{y})]_{2n+1}^{k_{2n+1}} \in \mathbb{R},$$

where we write $f^i(\underline{x}) = ([f^i(\underline{x})]_1, \dots, [f^i(\underline{x})]_{2n+1})$. To show that the derivative $D_1 G : \mathbb{R}^{N+(2n+1)} \to \mathbb{R}^{2n+1}$ has maximal rank (i.e., rank $2n+1$) it suffices to show that for distinct $\underline{x}$ and $\underline{y}$ the set of all partial derivatives $\partial_{\underline{a}} G(\cdot, \underline{x}, \underline{y})$ spans $\mathbb{R}^{2n+1}$. Assume for a contradiction this is not the case, and that these vectors lie in a lower dimensional subspace. By taking linear combinations we conclude that for any family of polynomials $F_1, \dots, F_{2n+1}; \mathbb{R}^{2n+1} \to \mathbb{R}$ (of degree d) the vectors

$$(F_i(\underline{x}) - F_i(\underline{y}), F_i(f(\underline{x})) - F_i(f(\underline{y})) \dots, F_i(f^{2n}(\underline{x})) - F_i(f^{2n}(\underline{y}))$$

for $i = 1, \dots, 2n+1$, are linearly dependent, or equivalently, $\det(C) = 0$ where

$$C = \begin{pmatrix} F_1(\underline{x}) - F_1(\underline{y}) & F_1(f(\underline{x})) - F_1(f(\underline{y})) & \cdots & F_1(f^{2n}(\underline{x})) - F_1(f^{2n}(\underline{y})) \\ F_2(\underline{x}) - F_2(\underline{y}) & F_2(f(\underline{x})) - F_2(f(\underline{y})) & \cdots & F_2(f^{2n}(\underline{x})) - F_2(f^{2n}(\underline{y})) \\ \vdots & \cdots & \ddots & \vdots \\ F_{2n+1}(\underline{x}) - F_{2n+1}(\underline{y}) & F_{2n+1}(f(\underline{x})) - F_{2n+1}(f(\underline{y})) & \cdots & F_{2n+1}(f^{2n}(\underline{x})) - F_{2n+1}(f^{2n}(\underline{y})) \end{pmatrix}.$$

Moreover, we can write

$$C = \begin{pmatrix} F_1(\underline{x}) & \cdots & F_1(f^{2n}(\underline{x})) & F_1(\underline{y}) & \cdots & F_1(f^{2n}(\underline{y})) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ F_{2n+1}(\underline{x}) & \cdots & F_{2n+1}(f^{2n}(\underline{x})) & F_{2n+1}(\underline{y}) & \cdots & F_{2n+1}(f^{2n}(\underline{y})) \end{pmatrix}$$

$$\times \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix}. \tag{B.1}$$

Since, by hypothesis, neither $\underline{x}$ nor $\underline{y}$ is a periodic point of period at most $2n+1$, we can see that for appropriate choices of $F_1, \dots, F_{2n+1}$ the values of the functions at the points $\underline{x}, \dots, f^{2n}\underline{x}, \underline{y}, \dots, f^{2n}\underline{y}$ can be taken to be any arbitrary values. In particular, we can arrange for the first matrix in (B.1) to have full rank or, equivalently, the corresponding linear map is surjective. The second matrix in (B.1) is given explicitly, and the corresponding linear map can be seen to be surjective. From (B.1), we see that the matrix C has non-zero determinant, since the corresponding linear map is surjective (being a composition of two surjective maps). This gives the required contradiction and we see that the derivative $D_1G : \mathbb{R}^{N+(2n+1)} \to \mathbb{R}^{2n+1}$ has maximal rank (i.e., rank $2n+1$). By the Rank Lemma we then see (as in the proof of Theorem 1) that locally those $\underline{a}$ such that $G(\underline{a}, \underline{x}, \underline{y}) = 0$ lie on a countable union of codimension one manifolds $X_0 \subset \mathbb{R}^N$.

To show that $\Phi_{(f,g)} : M \to \mathbb{R}^{2n+1}$ is an immersion consider the map

$$H : \mathbb{R}^N \times TM \to \mathbb{R}^{2n+1}$$

given by $H(\underline{a}, x, v) = D_x\Phi_{(f,g)}(v)$, where $g : M \to \mathbb{R}^{2n+1}$ is the polynomial associated to a $\underline{a} \in \mathbb{R}^N$. As in the proof of Theorem 1, $\Phi_{(f,g)}$ is an immersion if we have that $\underline{0} \notin H(\underline{a}, SM)$. We can again use expression (A.1) for this map, but now the entries (A.2) are of a more complicated form. More precisely, if we write $H(\underline{a}, \underline{x}, v) = (H_1(\underline{a}, \underline{x}, v), \dots, H_{2n+1}(\underline{a}, \underline{x}, v))$ we can use the chain rule to write

$$H_i(\underline{a}), \underline{x}, v) = Dg(\underline{a}, f^i\underline{x}, \cdot)D(f^i)(\underline{x}, v) \tag{B.2}$$

for $i = 1, \dots, 2n+1$, where:

(a) the derivative of $g : \mathbb{R}^{2n+1} \to \mathbb{R}$ at $f^i(\underline{x})$ is

$$Dg(f^i\underline{x}) = \left(\frac{\partial g}{\partial x_1}(f^i\underline{x}), \dots, \frac{\partial g}{\partial x_{2n+1}}(f^i\underline{x})\right),$$

where $\frac{\partial g}{\partial x_j}(\underline{x}) = \sum_{1 \leqslant |\underline{k}| \leqslant d} (k_j a_{\underline{k}})\underline{x}^{\underline{k}^{(j)}}$, $j = 1, \dots, 2n+1$; and

(b) the derivative of $f^i : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ at $\underline{x}$ is

$$Df^i(\underline{x}, \cdot) = \begin{pmatrix} \dfrac{\partial [f^i(\underline{x})]_1}{\partial x_1} & \cdots & \dfrac{\partial [f^i(\underline{x})]_{2n+1}}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial [f^i(\underline{x})]_1}{\partial x_{2n+1}} & \cdots & \dfrac{\partial [f^i(\underline{x})]_{2n+1}}{\partial x_{2n+1}} \end{pmatrix},$$

where we write the ith iterate of the map $f : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ in terms of its components as $f^i(\underline{x}) = ([f^i(\underline{x})]_1, \dots, [f^i(\underline{x})]_{2n+1})$.

We can deduce from (a) that

$$\partial_{\underline{a}} D_g(f^i(\underline{x})) = (k_1[f^i(\underline{x})]^{\underline{k}^1}, k_2[f^i(\underline{x})]^{\underline{k}^2}, \dots, k_{2n+1}[f^i(\underline{x})]^{\underline{k}^{2n+1}}),$$

where $[f^i(\underline{x})]'^{\underline{k}^j}$ is the evaluation of the monomial $k_j\underline{x}^{\underline{k}^j}$ at the point $f^i(\underline{x}) = ([f^i(\underline{x})]_1, \dots, [f^i(\underline{x})]_{2n+1})$. Then we can deduce from (B.2) that

$$\begin{aligned} &\partial_{\underline{a}} H_1(\underline{a}, \underline{x}, v) \\ &= \begin{pmatrix} k_1[\underline{x}]^{\underline{k}^1} & \cdots & k_{2n+1}[\underline{x}]^{\underline{k}^{2n+1}} \\ \vdots & \ddots & \vdots \\ k_1[f^{2n}(\underline{x})]^{\underline{k}^1} & \cdots & k_{2n+1}f^{2n}([\underline{x}])^{\underline{k}^{2n+1}} \end{pmatrix} \begin{pmatrix} \dfrac{\partial [f^i(\underline{x})]_1}{\partial x_1} & \cdots & \dfrac{\partial [f^i(\underline{x})]_{2n+1}}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial [f^i(\underline{x})]_1}{\partial x_{2n+1}} & \cdots & \dfrac{\partial [f^i(\underline{x})]_{2n+1}}{\partial x_{2n+1}} \end{pmatrix} \\ &\quad \times \begin{pmatrix} v_1 \\ \vdots \\ v_{2n+1} \end{pmatrix}. \end{aligned}$$

We want to show that the span of $\partial_{\underline{a}} H_1(\underline{a}, \underline{x}, v)$ has dimension $2n+1$ as we range over all $\underline{a} \in \mathbb{R}^N$. Taking linear combinations of these expressions gives

$$\begin{pmatrix} F_0(\underline{x}) & \dots & F_{2n}(\underline{x}) \\ \vdots & \ddots & \vdots \\ F_0(f^{2n}\underline{x}) & \dots & F_{2n}(f^{2n}(\underline{x})) \end{pmatrix} \begin{pmatrix} \dfrac{\partial[f^i(\underline{x})]_1}{\partial x_1} & \cdots & \dfrac{\partial[f^i(\underline{x})]_{2n+1}}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial[f^i(\underline{x})]_1}{\partial x_{2n+1}} & \cdots & \dfrac{\partial[f^i(\underline{x})]_{2n+1}}{\partial x_{2n+1}} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_{2n+1} \end{pmatrix}, \tag{B.3}$$

where $F_0, \dots, F_{2n} : \mathbb{R}^{2n+1} \to \mathbb{R}$ are polynomials. Since by hypothesis $\underline{x}, \dots, f^{2n}(\underline{x})$ are distinct, we can choose the functions $F_0, \dots, F_{2n}$ so that the first matrix in (B.3) can be of any form required.

Assume that for a given $\underline{a}$ we have $\underline{x} \in M$ and $v \in T_{\underline{x}}M$ such that $H(\underline{a}, \underline{x}, v) = \underline{0}$. Since $D_1 H$ has maximal rank $2n+1$, we can conclude that DH has maximal rank. By the Rank Lemma, we see that locally those $\underline{a}$ such that $H(\underline{a}, \underline{x}, v) = \underline{0}$ must lie on a codimension one manifold $X_1 \subset \mathbb{R}^N$. This completes the proof. □

References

[1] K. Alligood, T. Sauer, J. Yorke, Chaos, An Introduction to Dynamical System, Springer, Berlin, 1997.
[2] J. Froyland, Introduction to Chaos and Coherence, Institute of Physics, Bristol, 1992.
[3] A. Kolmogorov, Théorie génerale des systèmes dynamiques et mécanique classique, Proceedings of the International Congress of Mathematicians, North-Holland, Amsterdam, 1957, pp. 315–333.
[4] J. Milnor, A Nobel prize for John Nash, Math. Intelligence 17 (1995) 11–17.
[5] J. Nash, Real algebraic manifolds, Ann. of Math. 56 (1952) 405–421.
[6] N. Packard, J. Crutchfield, D. Farmer, R. Shaw, Geometry from time series, Phys. Rev. Lett. 45 (1980) 712–716.
[7] T. Sauer, J. Yorke, M. Casdagli, Embedology, J. Statist. Phys. 65 (1991) 579–616.
[8] M. Spivak, A Comprehensive Introduction to Differential Geometry, Vol. I, Publish or Perish, Berkeley, 1979.
[9] F. Takens, On the Numerical Determination of the Dimension of an Attractor, Lecture Notes in Mathematics, Vol. 1125, Springer, Berlin, pp. 99–106.
[10] F. Vivaldi, Arithmetical theory of Anosov diffeomorphisms, Proc. Roy. Soc. London Ser. A 413 (1987) 97–107.
[11] H. Whitney, Differentiable manifolds, Ann. of Math. 37 (1936) 645–680.